Trade-offs between spatial and temporal resolutions in stochastic super-resolution microscopy techniques

Jean-François Rupprecht, Ariadna Martinez Marrades, Gilles Tessier

▶ **To cite this version:**

Jean-François Rupprecht, Ariadna Martinez Marrades, Gilles Tessier. Trade-offs between spatial and temporal resolutions in stochastic super-resolution microscopy techniques. 2016. hal-01288853

HAL Id: hal-01288853
https://hal.science/hal-01288853

Preprint submitted on 22 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Trade-offs between spatial and temporal resolutions in stochastic super-resolution microscopy techniques

Jean-François Rupprecht,[1] Ariadna Martinez-Marrades,[2] and Gilles Tessier[2]

[1] *Mechanobiology Institute, National University of Singapore, 5A Engineering Drive 1, 117411 (Singapore).**

[2] *Sorbonne Universités, UPMC Univ Paris 05, Paris (France).*

(Dated: March 21, 2016)

Widefield stochastic microscopy techniques, such as PALM or STORM, rely on the progressive accumulation of a large number of frames, each containing a scarce number of super-resolved point images. We justify that the redundancy in the localization of detected events imposes a specific limit on the temporal resolution. Based on a theoretical model, we derive analytical predictions for the minimal time required to obtain a reliable image at a given spatial resolution, called image completion time. In contrast to standard assumptions, we find that the image completion time scales logarithmically with the ratio of the image size by the spatial resolution volume. We justify that this non-linear relation is the hallmark of a random coverage problem. We propose a method to estimate the risk that the image reconstruction is not complete, which we apply to an experimental data set. Our results provide a theoretical framework to quantify the pattern detection efficiency and to optimize the trade-off between image coverage and acquisition time, with applications to 1, 2 or 3 dimension structural imaging.

INTRODUCTION

Optical microscopy is a convenient tool to study biological processes, but its resolution is fundamentally limited by Abbe's diffraction. The image of a point source is a pattern whose size is comparable to the optical wavelength ($\sim 250\,\mathrm{nm}$), hence source points separated by a distance smaller than a wavelength are hardly distinguishable [1]. Electron microscopy provides a higher spatial resolution ($\sim 1\,\mathrm{nm}$) but at the cost of a more complex sample preparation which is incompatible with *in vivo*-imaging [2]. The recently developed super-resolution imaging techniques aim at combining the best of these two worlds. Using these techniques, spatial resolution as low as 10 nm have been achieved for imaging biological cell structures. However, their applicability to the study of dynamical biological processes is limited by their long acquisition times [3, 4].

Though relying on different optical probes, the super-resolution techniques known as PALM (Photoactivation Localization Microscopy) or STORM (Stochastic Optical Reconstruction Microscopy) rely on a common principle: sources that lie within the same diffraction-limited volume are separated by a sequential activation process, which introduces a temporal separation between source points [5]. Within each frame, a small and random fraction of probes is activated by illumination. This sparse subset of randomly activated probes is imaged to produce a frame. Then, finding the centroid of each diffraction patterns leads to a set of coordinates, having a nanometer-level precision [1, 6]. Merging all the single-molecule positions obtained on successive frames produces the final image. Since only a small fraction of probes is imaged per frame, a certain number of frames is required in order to obtain a reliable image. Multiplying this number by the typical acquisition time of frames (typically in the $10-100$ms range), we obtain the minimal time, denoted T, to obtain an image at a nanometer-scale resolution. A typical reported value is $T \sim 30\,\mathrm{min}$ for a whole cell imaging at a $\sim 10\,\mathrm{nm}$ resolution [4]. This value is too large to study many dynamical processes that occur in living cells, such as the contraction of acto-myosin units [7], reorganization of focal adhesion complexes [8] or protein cluster formation within the plasma membrane [9, 10]. Efforts to reduce the acquisition time are hindered by the risk that the collected set of observations is incomplete.

In this article, we are interested in the risk of *stochastic aberration* in the final reconstructed image, whereby an incomplete observation would incorrectly suggest the presence of gaps within the structure. An illustrative example is the case of 1D structures, such as DNA segments [12] or actin filaments [11] (see Fig. 1). Suppose that the reconstructed super-resolved image contains a hole: how can we reliably discriminate whether this hole is a genuine gap in the structure rather than an aberration due to a lack of observations?

This question illustrates the interest of our model, which provides an estimate for the minimal number of observation frames required to minimize the risk of stochastic aberrations. It also highlights a major difference between stochastic and deterministic imaging techniques. With raster-scan-based techniques, e.g. STimulated Emission Depletion (STED), the resolution is exempt of stochastic aberration but the acquisition time t increases linearly with the size of the field of view [13]. With widefield stochastic techniques such as PALM, the imaging time is thought to be independent of the size of the field of view. The density of the observed points, and concomitantly the final image resolution, increase with the acquisition time t. Image size, density and total

*Electronic address: mbijr@nus.edu.sg

acquisition time are therefore essential parameters to assess the relevance and performance of widefield stochatic techniques.

We briefly outline a simple reasoning that leads to the incorrect conclusion that the image completion time should be independent of the size of the field of view. We first suppose that fluorescent events are distributed according to a homogeneous Poisson process, such that the probability density dP that an event occurs in an infinitesimal space of volume ds reads $dP = \rho ds$ [14]. We now consider a regular domain of volume S within a D dimensional space, in which we assume a constant density of fluorophore d. Furthermore, we assume that at each frame, only a fraction f of fluorophores are detected. The number of detected fluorescence events after one frame, denoted $N^{(1)}$, is a Poisson process of density $\rho = fd$; hence $P\left[N^{(1)} = n\right] = \exp(-\rho S)(\rho S)^n/n!$. If $S = A$ refers to the volume of the Abel diffraction pattern, the mean number of fluorescence events per frame ρA should be lower than 1 in order to limit the risk of overlapping point spread functions. Typically $A = (10^2\,\mathrm{nm})^D$ [5], hence $\rho < 10^{-2D}\mathrm{nm}^{-D}$. In the case of membrane ($D = 2$) with fluorophore density $d = 10^4\,\mu\mathrm{m}^{-2}$, the corresponding maximal fraction of activated fluorophores should be $f < 10^{-3}$. Then, after a number T of frames, the total number of collected events is distributed according to a Poisson distribution, with $P\left[N^{(T)} = n\right] = \exp(-\rho ST)(\rho ST)^n/n!$. This number of collected events determines the resolution of the reconstructed image, due to the Nyquist criterion, which states that the mean distance between sampling points must be at least two-fold smaller than the desired resolution [15–17]. Therefore, to attain a resolution of $\delta s = (10\mathrm{nm})^D$, a minimal requirement is to collect at least 2^D evenly spaced observations per resolution volume δs, or, equivalently, to have at least 1 observation per volume $\sigma = 2^{-D}\delta s$. As the density of detected fluorophores up to the time T reads ρT, the Nyquist criterion leads to the identity $\rho T\sigma = 1$, in order to guarantee that, on average, there has one observation per elementary volume σ. With $\sigma = 2^{-D}\delta s$ and $\rho = 10^{-2D}\mathrm{nm}^{-D}$, this leads to a minimal number of frames $T = 2^D \cdot 10^D$. In 2D and with a typical frame rate of 10 Hz, the corresponding acquisition time would be of the order of a few minutes, independently of the size of the image. The relation $T\sigma = \rho^{-1}$ also illustrates how recent imaging techniques achieved live-cell imaging at the cost of a lower spatial resolution [16–20]. However, the condition $\rho T\sigma = 1$ does not imply that every elementary volumes have collected their one observation: some elementary volume may have collected several observations while most others none.

In this paper, we argue that, due to the randomness in the localization of events in stochastic widefield microscopy, the imaging time should not be considered independent of the size of the image. The uneven spatial distribution of events leads to a correction term $\ln(S/\sigma)$ in the expression of the imaging time. Hence the main result of our article lies in the following relation:

$$T\sigma = \rho^{-1} \quad \underset{\text{this article}}{\longrightarrow} \quad T\sigma = \rho^{-1}\ln(S/(\sigma\theta)) \qquad (1)$$

which means that the trade-off between the spatial (σ) and temporal (T) resolutions is not only regulated by the density of activated fluorophores per frame ρ, but also by the ratio of the size of the field of view S by the desired spatial resolution, and by an additional parameter $\theta < 1$ which characterizes the reliability of the final image.

In many situations of interest, the prefactor $\ln(S/\sigma)$ cannot be assumed to be close to 1. For example, a cell of extension $S = 10^3\,\mu\mathrm{m}^2$ contains 10^7 squares of area $\sigma = 10^{-4}\,\mu\mathrm{m}^2$ (ie. a typical size for an Abel diffraction pattern). This corresponds to the ratio $\ln(S/\sigma\theta) = 16$. Any specific choice for the value of θ only increases the value of the acquisition time. Hence our result predicts that the image completion time should be one order of magnitude longer than what would be expected from the arguments leading to expression in the left hand side of Eq. (1).

The paper is organized as follows. We first present the experimental setup that we use to test the validity of our theoretical expressions, in which we measure the light scattered by Brownian nano-particles at the surface of a two-dimensional sample (Method – Experiments). We then define two image rendering schemes (Method – Image rendering). We then prove the relation Eq. (1), and we show its connexion to the coupon-collector problem [29–31]. Therefore, we refer to the result of Eq. (1) as the *coupon-collector scaling*.

We then consider the robustness of the coupon-collector scaling for several requirement on the image completion. We first consider the effect of correlations between successive frames. This case is particularly motivated by our experimental setup, in which the escapes and returns of the Brownian particles within the detection zone leads to temporally correlated scattering events between successive frames. We point out that there is a close analogy between our experiments and PALM techniques in which the fluorophores blinking statistics exhibit time-correlations [28]. We recall that bleaching refers to an irreversible transition of a probe to an inactive state, which can occurr following a fluorescence event [5]. The analogy holds between blinking events, identified to light-scattering events from Brownian particles, and bleaching events, identified to the escape events far from the illuminated region and with no ulterior return.

We show that the coupon-collector scaling does not hold when a large number of coverage is required. We then compare our theoretical predictions to our experimental results. We conclude our article by presenting a procedure to estimate the risk of stochastic aberration, in real-time during the acquisition (ie. at each frames during the image acquisition process).

METHODS

Experiments

We briefly present the experimental data that we will use to test our predictions on the coverage time. In our recent work [27], we present a new stochastic imaging technique to map an electromagnetic field with a nanoscale resolution using light-scattering Brownian particles as local probes of the intensity field. The Brownian motion of the scatterers eventually lead to a full coverage of the imaged field. Following [27], we consider the imaging problem of an evanescent wave created by a Total Internal Reflection Microscopy setup. In this setup, we consider that the optical intensity field can be modelled as

$$I(x, y, z) = I_0(x, y) \exp(-z/\beta(x, y)), \tag{2}$$

where β is the penetration length of the field, and I_0 is proportional to the optical intensity of the field at the surface – with a proportionality constant related to the scattering cross section of the particles. In principle, in most situations, both quantities β and I_0 can vary with the location (x, y) on the surface. In this context, the term image acquisition refers to the determination of the fields I_0 and β.

However, as a first test of the method, the experimental data set from [27] corresponds to a situation where both I_0 and β are homogeneous within the whole field of view. We detail a procedure that leads to the determination of I_0 and β in the SI, C.

Two image rendering methods

Super-resolution techniques rely on the localization of the center of diffraction spots, which provides a set of points. However, a spatial extention needs to be attributed to each point to obtain an image that is readable to the human eye. In the following, we will be interested in the two following image rendering methods: (i) the BFM (BFM), which is adapted to a density image representation [21] and (ii) the patch method (PM), which is associated to a pointillist representation [22].

The BFM considers the structure of interest as tessellated into F square pixels of equal area, which can therefore be expressed as the ratio of the total volume by the resolution volume: $F = S/\sigma$. Each new event falls within a specific pixel, thereby increasing by one the cumulative number of observations of this pixel. This method is naturally adapted to measure the densities. Though we employ the term pixel in the following, our method also applies to 3D imaging problems in which F refers to the number of voxels within the structure of interest [23–26].

In the other hand, the PM associates to each event a surrounding extension, characterized by the quantity σ, which is either a length (1D), an area (2D) or a volume (3D). Generally, the spatial extension is chosen to correspond to the spatial uncertainty associated to the localization procedure (e.g. a few nanometers, [22]). The image completion time is related to the minimal number of patches required to cover the structure of interest.

Statistics of events

As discussed in the Introduction, we consider that the number of events occurring on each frame follows a Poisson statistics, with a mean $\mu = \rho S$, where S is the volume of the region of interest, and ρ is the density of events per frame. We consider that the density of events ρ (i) is homogeneous within the volume of the structure of interest (ii) is time-independent.

Concerning the homogeneous assumption, we point out that, in most cases of interest, the field of view cannot be considered as homogeneous. Indeed, the field of view generally contains regions where no events should occurr. For example, consider the localization problem of actin filaments within a biological cell by PALM. Points which are located outside of the actin filaments should not yield any observation. We assume that there is a one to one correspondence between pixels which contain fluorophores and pixels within the structure. This amounts to neglecting the background noise, ie. spurious detection events in regions where fluorophores should not be present. Hence, we are interested in the imaging process of a subset of the field of view, which we call the structure of interest. Similarly, in the context of the experimental application of our method with Brownian nanoparticles, we expect that only a subset of the field of view is relevant, as characterized by significantly higher intensity fields. This corresponds to the experimentally relevant cases of either structured [39] or rough [40] metallic surfaces, which can exhibit local hotspots whereby the incident light is concentrated on the nanometre scale to produce an intense electromagnetic field. This is particularly justified as we expect the probability to observe a particle to be a function of both β and I_0. In both experimental contexts, we assume that the event probability follows a homogeneous Poisson process within the volume of the structure of interest

In PALM, the time-independent assumption can be questionned due to bleaching events [28]. However, these bleaching events have no effect on the image completion time neither in the PM representation, nor in the BFM configuration with $r = 1$. In the BFM with $r > 1$, the time-independent assumption is justified if the number of the remaining unbleached fluorophores per pixel is sufficiently large.

Estimation of the structure size

We emphasize that, in most situations, the volume of the structure is unknown *a priori*. Within the BFM, we show in the SI A 1 that the maximum likelihood estimator

of the number of relevant pixels F corresponds to the quantity:

$$\widehat{F}^{(t)} = \sum_{j=1}^{F} \min\left(M_j^{(t)}, 1\right), \tag{3}$$

where $M_j^{(t)}$ is the cumulative number of measures of the pixel j, e.g. $M_j^{(t)} = 0$ if the pixel j has never collected any event up to time t and $M_j^{(t)} \geq 1$ if the pixel has been observed at least once up to time t (see Fig. 1). Similarly, within the patch-method framework, the maximum likelihood estimator of the structure volume consists in the covered volume at the time t. These two estimators are biased, as they tend to underestimate the structure volume.

Mathematical definition of the image completion time

We call *image completion time* the minimal number of frames required to obtain a *complete* image of the structure of interest. The term complete refers to the condition that every pixel or point (among those that should be observed) has been covered at least a certain number of times, denoted $r \geq 1$. More precisely, the image completion time T is the random variable (called stopping time) that corresponds to the minimal time t such that $\min_j \left(M_j^{(t)}\right) = r$; where $j \in [1, \dots F]$ in the BFM framework, or j refers to any point within the volume of interest in the PM framework. We will be mainly interested in the centile of T, denoted t_θ and defined as:

$$\mathbb{P}\left[T \leq t_\theta\right] = \mathbb{P}\left[\min_j \left(M_j^{(t_\theta)}\right) \geq r\right] = 1 - \theta, \tag{4}$$

where θ is the tolerated risk tolerance level. To summarize, the quantity $t_{0.05}$ refers to the minimal number of frames that guarantees, with 95% probability, that the image is complete.

Simulations

Both in the BFM and PM frameworks, the volume of the structure of interest is tessellated into a grid of elementary squares. In the BFM, each event covers a single elementary square; while in the PM, each patch σ covers a square matrix of elementary squares. In both frameworks, we generate a large sample of coverage events and we analyse the resulting distribution of coverage times using Matlab's *prctile* function.

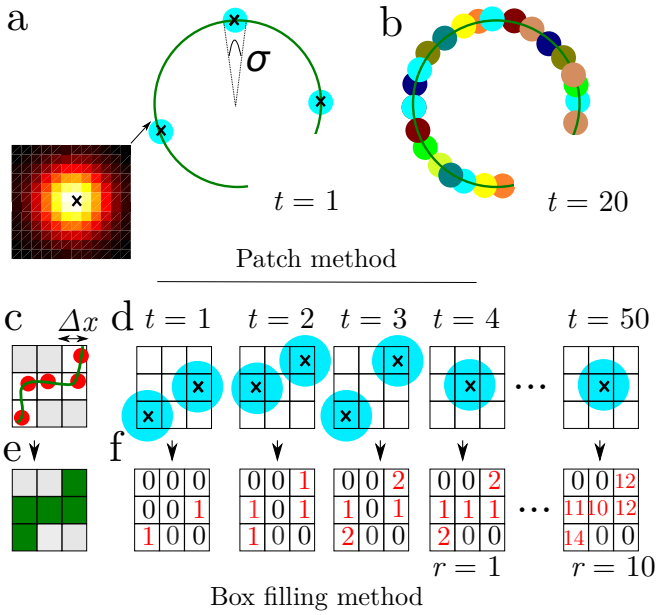

Figure 1: (Color online) Structural reconstruction by stochastic super-localization microscopy. Probes (colored dots) are bound to a structure of interest (green line). (a–b) Circular patches representation: (a) Upper left inset: Abel diffraction pattern observed in a CCD camera. The super-resolution algorithm yields a set of coordinate corresponding to the center of the pattern (black cross). In the patch method representation, each point coordinate is represented by a disk with a radius σ that is proportional to the uncertainty of the super-localization procedure (blue disk). (b) Patches accumulate with the acquisition time, eventually covering the whole structure of interest (patches are represented by different colors patches for separate frames). (c–f) Box-filling representation, which leads to a density map in terms of a number of accumulated events per pixel. (c) The field of view is divided into $N = 9$ pixels among which $F = 5$ pixels contain probes. (d) A sequence of frames (blue circle: size of the Abel pattern). (e) Target image. (f) Map of the cumulative number of observations $M_j^{(t)}$, for each j and for each frame t. Complete image completion (with $r = 1$) is obtained after $t = 4$ frames. At $t = 50$, all pixels have been observed at least $r = 10$ times.

RESULTS AND DISCUSSION

The image completion time follows a coupon-collector scaling

We first derive the main result of Eq. (1) in the case of the BFM with no time-correlation between frames. We first suppose that F and probability p_1 are known quantities. Under the assumption that detection events occurring in separate pixels are independent, the probability that exactly M pixels have been observed at least once ($r = 1$) reads:

$$\mathbb{P}\left[\widehat{F}^{(t)} = M\right] = \binom{F}{M} p_0^{(M-F)t} \left(1 - p_0^t\right)^M, \tag{5}$$

where $p_0 = 1 - p_1$, and $p_1 = \rho\sigma$ is the probability that an event occurs in a given pixel and at a given frame. Equivalently, we can express this probability as $p_1 = \mu/F$. In particular, we conclude that the probability that the estimator $\widehat{F}^{(t)}$ is equal to its target value F is equal to:

$$\mathbb{P}\left[\widehat{F}^{(t)} = F\right] = \left(1 - (1-p_1)^t\right)^F . \tag{6}$$

As expected, the value of the probability in Eq. (6) converges to 1 with time (see Fig. 5c).

We now determine the centile of the image completion time, defined according to Eq. (4) as the solution of the equation $\mathbb{P}\left[\widehat{F}^{(t_\theta)} = F\right] = 1 - \theta$, hence $t_\theta = \ln\left(1 - (1-\theta)^{1/F}\right) / \ln\left(1 - p_1\right)$. In the limit $p_1 \ll 1$ and for sufficiently high centiles ($\theta < 0.1$), we find that the centile of the imaging time reads:

$$t_\theta \underset{1 \ll F}{\sim} \frac{F}{\mu} \ln\left(\frac{F}{\theta}\right) = \frac{1}{\rho\sigma} \ln\left(\frac{S}{\theta\sigma}\right), \tag{7}$$

where we use the identities $p_1 = \mu/F$, $p_1 = \rho\sigma$, and $F = S/\sigma$. The latter expression corresponds to the announced expression in Eq. (1).

The key feature of Eq. (7) is the non-linear dependence of the imaging time in terms of the number F of pixels that characterize the structure. This scaling is related to the classical coupon-collector problem [29–31]. The problem consists in buying a minimal number of the boxes (each containing a random coupon) in order to gather a complete collection of coupons, with a sufficiently high probability. Here, we focus on the case where each box contains, at random, either 0 (with probability p_0) or 1 coupon – in which case the mean number of coupons per box is equal to $\mu = 1 - p_0$. A straightforward proof leads to the following exact expression for the mean number of bought boxes t (i.e. frames) required to collect all coupons (i.e. all pixels) is $\mathbb{E}[T] = F(1 + 1/2 + \ldots + 1/F)/\mu$. If the number of coupons F is large, the latter expression takes the asymptotic form $\mathbb{E}[T] = F\ln(F)/\mu$. Adapting the identity (2) of Ref. [29], one shows that the centile of the stopping time reads $t_\theta = (F/\mu) \times \ln(F/\theta)$ in the same limit $F \gg 1$, which corresponds to Eq. (6) after identification of the mean number of coupons per box to the mean number of events per frame. The coupon-collector analogy also highlights the fact that the number of frames required to gather the very last remaining pixel is very long, as it scales linearly with the total number of pixels $F \gg 1$.

Mind that Eq. (7) weakly depends on the risk level θ, which is another characteristic property of the coupon-collector problem [29–31]. A heuristic justification is that, the stopping times distribution shrinks to its mean as the number of pixels F is increased.

Furthermore, the image completion time does not depend on the number of pixels, but not on the 1D, 2D or 3D nature of the structure. This is expected since pixels are considered to be independent.

Finally, we present a heuristic justification for the coupon-collector scaling, based on the evolution of the mean number of observed pixels. From Eq. (5), one easily finds that the mean number of observed pixels reads $\langle\widehat{F}\rangle = F\left(1 - p_0^t\right)$, which, in the limit $\mu/F \ll 1$, becomes $\langle\widehat{F}\rangle = F\left(1 - \exp(-\mu t/F)\right)$. We then notice that the time needed for the number of observed pixels to be $\langle\widehat{F}\rangle = F(1-\theta)$ follows the coupon-collector scaling of Eq. (7).

The coupon-collector scaling holds when redundant observations per pixel are required

We show that the coupon-collector scaling holds if the image completion is reached when every pixels have accumulated at least a number $r > 1$ of observations. We show in the SI A 3 a that, in this case, the centile of the image completion time reads

$$t_\theta \underset{1 \ll F}{\sim} \frac{F}{\mu} \left\{\ln\left(F/\theta\right) + (r-1)\ln\left(\ln F\right)\right\}. \tag{8}$$

This relation corresponds to the centile of the coupon collector's problem when r copies of each coupon need to be collected (see [29, 34]). We emphasize that Eq. (8) requires that the number of coverage is sufficiently small, ie. that $r \ll \ln(F)$.

We also generalize the result of Eq. (8) to a multi-color imaging problem (see SI A 2).

The coupon-collector scaling holds with the patch image-rendering method

We now consider that the image results form the accumulation of circular patches, whose radius σ corresponds to the spatial resolution. The patch centers are distributed according to a homogeneous Poisson distribution [45] within the structure of interest, of volume S.

The study of coverage problem has a long history [35, 36]. However, analytical results concerning coverage problems in two dimensions are rather recent [37, 38]. These studies were motivated by the study of the wifi coverage resulting from randomly located routers. We will make use of results concerning the expression of the centile n_θ of the number of patches required to cover a circle [35] or a square [37] by circular patches.

Here, we seek an expression of the centile time t_θ, i.e. a time expressed in terms of a number of frames t, rather than the centile time expressed in terms of the number of patches n. We expect that $t_\theta = n_\theta/\mu$ where μ is the number of events per frame. Indeed, in the small patch limit $\sigma/S \ll 1$, full coverage events occur when the number of events is large ($n \gg 1$) in which case the number of events is simply proportional to the number of frames t. This approximation is further justified in the

SI. Therefore, following Refs. [35] and [37], we expect the centile time t_θ to obtain a r-fold coverage of a D-dimensional volume S by circular patches of volume σ to read:

$$t_\theta \underset{\sigma/S\ll 1}{\sim} \frac{S}{\mu\sigma}\left\{\ln\left(\frac{S}{\sigma\theta}\right)+(r+\gamma_D)\ln\left[\ln\left(\frac{S}{\sigma}\right)\right]\right\}, \quad (9)$$

where $\gamma_1 = 0$ in 1D [35] and $\gamma_2 = 2$ in 2D [37]. It appears that no analytical expression exist concerning the 3D coverage problem. Our simulations for the coverage problem by squares (defined in Method – Simulations) agree with the scaling defined Eq. (9) and with the value of the constants γ (see SI, Fig. 6). In 3D, we find that $\gamma_3 = 3$.

Remarkably, Eq. (9) takes a similar form as the coupon-collector problem from Eq. (8). This similarity suggests that in the limit $\sigma/S \ll 1$, regularly spaced patches of size σ/S behave as if they were independent. Mind, however, that the expression from Eq. (8) corresponds to a value $\gamma_D = -1$ for any space dimension. The origin of this discrepancy at second order in the ratio $\sigma/S \ll 1$ is discussed in Ref. [35].

In conclusion, we have shown that both the path and box-filling image rendering methods lead to similar expressions for the image completion time, which are characterized by a logarithmic dependence in terms of the ratio of the image volume by the spatial resolution volume. Mind that, in the BFM, the latter ratio is equal to the number of pixels. Last, we point out that both Eqs. 8 and 9 hold provided that $r < \ln(S/\sigma)$.

The coupon-collector scaling holds in the presence of correlations between frames

In our experiments, the scatterer may enter, escape or return within the field of view, leading to correlated observations between successive frames. In contrast to the discussion leading to Eq. (6), these temporal correlations invalidate the independence hypothesis that allows to factorize the final time probability distribution. We encompass these correlated observations through the following box-filling model, in which the number of events per pixel and per frame is assumed to be a random variable K with a general probability law $p_k = \mathbb{P}(K = k)$ for all $k \geq 0$. The statistics of K encompass the effect of time-correlated observations by neglecting the time between successive correlated events. Comparison of this model to experiments is satisfactory, as visible in Figs. 2(e) and (f), in which we represent the experimental data from Ref. [27] and simulated evolutions of the cumulative number of events $M_j^{(t)}$.

In the following, we will make use of the average jump size ν and variance of the jump size σ^2, defined as $\nu = \sum_{k=1}^{\infty} kp_k$ and $\sigma^2 = \sum_{k=1}^{\infty} k^2 p_k - \nu^2$, respectively. We assume that the set of probabilities p_k, $k \geq 0$ is identical for each of the pixels of the structure to be imaged. The statistics of the jump distribution has a drastic impact

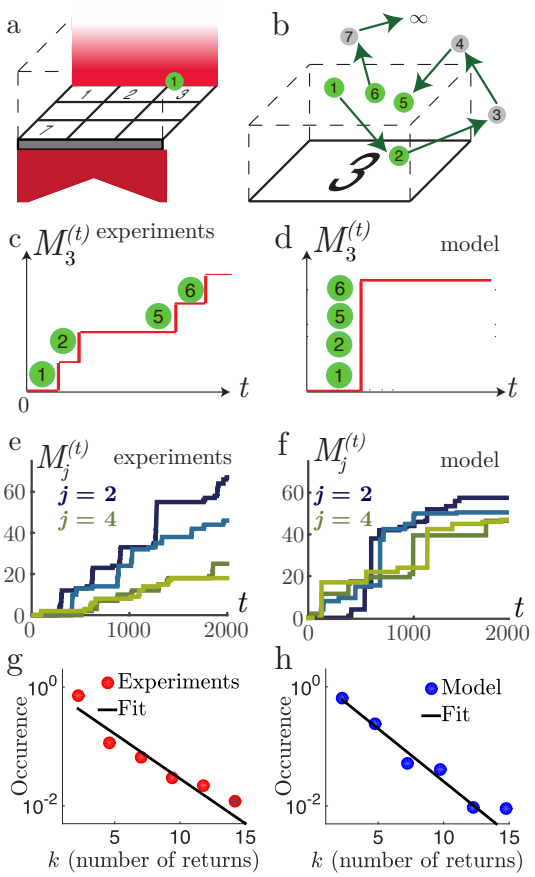

Figure 2: (Color online) Schematic view of the stochastic Brownian microscopy technique and its modeling. (a) Global view of the total field of observation, divided into F pixels (BFM framework). (b) Zoom on pixel 3, including a particle labelled by the frame instant. After $t \geq 6$, the probe is not detected again. (c-d) Scheme of the evolution of the cumulative number of observations $M_3^{(t)}$ occuring in pixel 3 (c) as seen experimentally (d) as represented in our model, where all correlated observations are collapsed into one single instantaneous events. (e-f) Evolution of the cumulative number of observations obtained, either (e) from the experimental data set of Ref [27], or (f) from Monte-Carlo simulations, with the fitted jump distributed. The different colors correspond to different pixels (4 among $F = 100$). (g-h) Distribution of jump length $p_k \propto \exp(-k/k_c)$ (in log-scale): (g) from experiments (blue circles), where the maximum likelihood estimator of the exponential model [44] provides the value $k_c = 2.9 \pm 0.1$ (black line); (h) from simulations, with $k_c = 2.9$ and the same number of 2792 simulated events.

on the value of the imaging completion time. In SI A 3 a, we show that the imaging completion time reads

$$t_\theta \underset{\sigma\ll S}{\sim} \frac{\left\{\ln\left(\frac{S}{\sigma\theta}\right)+(r-1)\ln\left[\ln\left(\frac{S}{\sigma}\right)\right]\right\}}{1-p_0}, \quad (10)$$

provided that $p_1 \neq 0$. The difference with Eq. (7) is due to the term $1 - p_0$, which is determined by the precise statistics of K and which may significantly differ form the value of μ/S. Therefore, at a constant total mean

number of events per frame μ, an increase in the mean number of correlated events ν also increases the imaging time.

We conclude that, temporal correlations can significantly affect the value of the image completion time, yet without affecting the coupon-collector scaling of the image completion time.

Situations in which the coupon-collector scaling does not hold

We first consider the case in which the image is defined as complete as soon as $M < F$ different pixels have been acquired. We find that the probability defined in Eq. (5) is maximal after a number of frames $t_{\rm opt}(M) \sim M/\mu$ in the limit $\mu/F \ll 1$ and $M/F \ll 1$. Hence, the image completion time is be proportional to M, with no logarithmic dependence on the parameters.

Secondly, we consider the case in which a large number of observations per pixel is required ($r \gg F$). We show that the coupon collector scalings from Eqs. 8 and 9 does not hold in this limit. We first remark that, due to the central limit theorem, the number of observations collected in the pixel j eventually converges with t towards a Gaussian distribution: $M_j^{(t)} \sim \mathcal{N}\left(t\mu/F, t\sigma^2/F\right)$, where μ and σ^2 are defined as the mean and variance of the total number of events per frame on the whole field of view, respectively. Under the Gaussian assumption, we find that the probability distribution of the image completion time T reads:

$$\mathbb{P}\left[T \leq t\right] = 2^{-1/F}\left\{1 - \mathrm{erf}\left(\frac{r - \mu t/F}{\sqrt{2\sigma^2 t/F}}\right)\right\}^F, \qquad (11)$$

where $\mathrm{erf}(x) = \int_{-\infty}^{x} \mathrm{d}t\, \exp(-t^2)/\sqrt{\pi}$ is the error function [42]. In the limit of a large number of observations $r \gg \ln(F)$, the expansion of the error function around 0 provides the following approximate expression:

$$t_\theta \sim \frac{Fr}{\mu} + \sqrt{\frac{2r\sigma^2}{\mu}\log\left(\frac{F}{2\sqrt{2\pi}\theta}\right)}. \qquad (12)$$

in the limit $r \gg F/\mu$. The key feature form Eq. (12) is that the image completion time t_θ does not follow the coupon-collector scaling from Eq. (7). It rather corresponds the result in the left hand side expression from Eq. (1). Indeed, the effects of the localization randomness is all the more averaged out that the required redundancy per pixel is increased.

We have obtained analytical results for the image completion time problem in the two regimes $r \ll \ln(F)$ and $r \gg F/\mu$. We resort to numerical simulation to describe the intermediate regime $\ln(F) < r < F/\mu$.

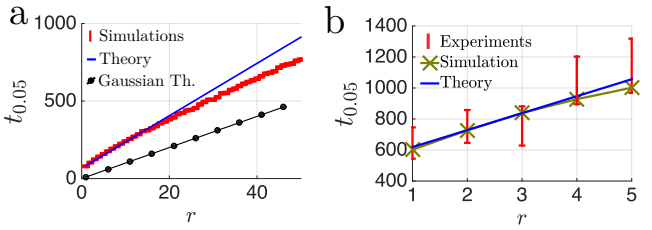

Figure 3: (Color online) Centile t_θ of the image completion time as a function of the required number of redundant observations per pixel r. (a) Simulations with at most one observation per pixel and per frame, $F = 15$ and $p_1 = 0.1$: (solid blue line) analytical expression from Eq. (8); (red error bars) centile estimation from stochastic simulations, with estimate incertitude; (black circle) approximate solution from Eq. (12). (b) Experiments from [27], with $F = 15$: (blue solid line) theoretical prediction from Eq. (8) with a jump probability distribution $p_k \propto \exp(-k/k_c)$ with $k_c = 2.9$; (green crosses) stochastic simulations; (red error bars) centile estimation from the experiments, with error bars obtained by bootstrapping [41].

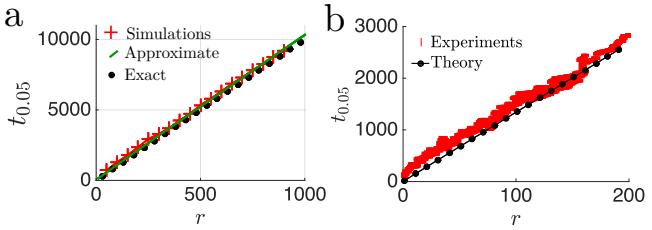

Figure 4: (Color online) Centile t_θ of the image completion time as a function of the required number of events per pixel in the regime $r \geq \ln(F)$. (a) Simulations with at most one observation per pixel, $F = 15$ and $p_1 = 0.1$: (solid blue line) exact solution of Eq. (11); (red line) centile estimation from stochastic simulations; (black circle) approximate solution from Eq. (12). (b) Experiments from [27], with $F = 4$: (red bars) centile estimation from experiments, where the error is estimated by bootstrapping [41]; (black dots) theoretical prediction from Eq. (11).

COMPARISON TO EXPERIMENTS

We represent the experiments data from Ref. [27] within the BFM framework and we include temporal correlations between frames. First, the mean number of particles per frame and over the whole field of view reads $\mu = 0.70$. Secondly, the jump distribution is estimated as follow: two successive events are assumed to correspond to the return of the same particle if (i) they occur within the same pixel and (ii) they are separated by a time interval of less than $\Delta = 5$ frames. The experimental histogram is fitted by the distribution $p_k = A_{k_c}\exp(-k/k_c)$, where $A_{k_c} = 1/(1 - \exp(-1/k_c))$, and $k_c = 2.9$ (see Fig. 2.g.). This leads to a mean jump size $\nu = 1/(1 - \exp(-1/k_c)) = 3.4$ and a variance $\sigma^2 = 1/(\cosh(1/k_c) - 1) = 17$. We check that our results weakly depend on the specific value attributed to

the separation time Δ.

As described in Figs. 3a and 4b, we show that our theoretical expressions from Eqs. (10) and (12) both fit to the experimental estimation of the centile time in their respective validity range. We point out that a straightforward implementation of Eq. (8), which would neglect temporal correlations, leads to a value that is an order of magnitude lower than what is experimentally observed.

In SI, C, we justify that a large number of redundant observations $r \approx 4 \cdot 10^3$ is required to overcome the statistical noise in the holographic height measure and to obtain a reliable measure of the length β.

REAL-TIME ESTIMATION OF THE RISK OF STOCHASTIC ABERRATION

Experimentally, the two quantities F and μ are unknown *a priori*. These quantities are indeed associated to the structure to be imaged, whose properties are unknown prior to imaging. Here, we propose a real-time procedure to determine whether we can safely consider that the image is complete. We emphasize that this procedure is not specific to a choice of image representation method, nor on the required number of redundant observations r.

After t frames, $\widehat{F}^{(t)}$ refers to an estimator of F; $\widehat{\mu}^{(t)}$ is an estimator of the mean number of events per frame. We represent the convergence with time of the estimator $\widehat{F}^{(t)}$ on Fig. 5.a.

Based on these estimators, we can estimate the probability that the image is complete. For example, in the box-filling representation with $r = 1$, the estimator of the image completion probability reads:

$$\widehat{\mathbb{P}}\left[\widehat{F}^{(t)} = F\right] = \left(1 - (1 - \widehat{\mu}^{(t)}/\widehat{F}^{(t)})^t\right)^{\widehat{F}^{(t)}}. \qquad (13)$$

We represent the evolution of the estimated probability corresponding to Eq. (13) in Fig. 5c. We set the values to $F = 100$ and $p_1 = 10^{-2}$. At $t = 300$ the image completion probability is lower than 4.10^{-3}: hence more frames are needed. The question is now to determine how many additional number of frames are required.

We use the expression for the image completion time to determine the required additional number of frames. Within the BFM with $r = 1$, an estimator of t_θ reads:

$$\widehat{t_\theta}^{(t)} = \frac{\widehat{F}^{(t)}}{\widehat{\mu}^{(t)}} \ln(\widehat{F}^{(t)}/\theta). \qquad (14)$$

We represent the convergence of Eq. (14) to the expected value of the centile time t_θ in Fig.1b,d. After $t = 300$ frames, we estimate that about 420 additional frames are required, which is consistent with the theoretical value of the centile time $t_\theta = 760$.

Based on the estimators for the probability that the image is complete and on the estimator of the image completion time, we propose the following procedure to analyse an imaging experiment in which t frames have been collected:

1. Compute the estimators of the number of pixels ($\widehat{F}^{(t)}$) and of the mean number of events per frame ($\widehat{\mu}^{(t)}$).

2. Compute the estimator of the probability that the image is complete. If this estimator is higher than a desired confidence threshold, the imaging process can be stopped.

3. Otherwise, compute the estimated image completion time $\widehat{t_\theta}^{(t)}$. Perform $\widehat{t_\theta}^{(t)} - t$ additional frames and return to step 1 with the substitution $t \leftarrow \widehat{t_\theta}^{(t)}$.

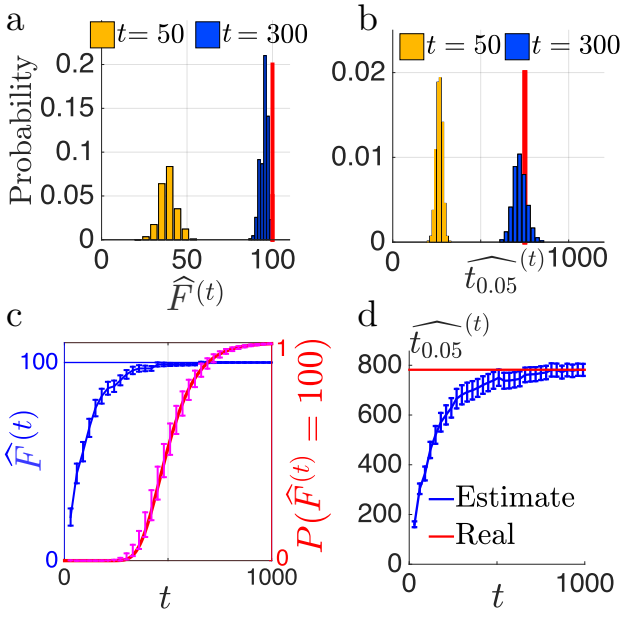

Figure 5: (Color online) Numerical simulation of the imaging method (see Eqs. (13– 14)). Here, the total number of pixels is $F = 100$, the probability for an event observation is $p_1 = 10^{-2}$ per frame and per pixel, and quantities are averaged over 10^4 samples. (a) Histogram of the values of the estimator $\widehat{F}^{(t)}$ after a number $t = 50$ of observations (light orange) and $t = 300$ (dark blue); (red vertical line) the limit value is $F = 100$. (b) Histogram of the values of the estimator $\widehat{t_{0.05}}^{(t)}$ after a number $t = 50$ of observations (light orange) and $t = 300$; (red vertical line) the limit value is $t_{0.95} = 760$. (c) Evolution of the mean value of the estimator $\widehat{F}^{(t)}$ (solid blue curve), together. We also represent the probability $P(\widehat{F}^{(t)} = F)$ for the image completion (solid magenta curve). In both cases, error bars indicate the standard deviation estimated from the random sampling. Hence a single random realization (i.e. a single experiment) is sufficient to obtain a good estimate of $\widehat{F}^{(t)}$ and $P(\widehat{F}^{(t)} = F)$. (d) Probability distribution for the estimator $\widehat{t_{0.05}}^{(t)}$. The distribution converges to the centile $t_{0.95} = 760$ (red vertical line) as t increases. The value $\widehat{t_{0.95}}^{(50)} = 264$ is significantly larger than the current number of frame $t = 50$: this is consistent with the conclusion that more observations are required.

We emphasize that this procedure is not specific to any particular criteria for the image completion. For example, in a situation where one expects that a large redundancy is required ($r \gg \ln(F)$), one can use the expressions of Eq. (11) for the probability that the image is complete and Eq. (12) for the image completion time. In the SI B 3, we provide an expression for the probability that the image is complete within the PM framework.

CONCLUSION

Our theoretical model provides a unified framework to describe the temporal resolution of several types of stochastic microscopy techniques. These include PALM, in which a large number of fluorescent probes are attached to the sample and are stochastically activated, or techniques in which a smaller number of scattering probes stochastically explore the imaged region. We derive analytical expressions for the centile of the imaging time with specific requirements for the completion of the imaging process. When only one or a few events are required per pixel, the temporal resolution is shown to be non-linearly coupled to the spatial resolution (pixel size), due to the spatial redundancy of detection events. However, the temporal resolution becomes linearly coupled to the spatial resolution when a large spatial redundancy of events is needed, as the effects of the localization randomness are averaged out. Our result for the imaging time are readily applicable to estimate the minimal time required to observe spatial patterns by stochastic imaging, with applications ranging from the detection of protein clusters by PALM [9] to the detection of the electromagnetic field around nano-antennas by Brownian particles [27].

SUPPORTING MATERIAL

Three figures and detailed proofs are available at XXX.

AUTHOR CONTRIBUTIONS

A. M. M. and G. T. designed the experiments, instigated the theoretical problem, carried out the experiments and data analysis, and commented on the manuscript; J.-F. R. performed the theoretical calculations, simulations, centile time analysis of the experiments and wrote the manuscript.

ACKNOWLEDGEMENTS

We thank V. Studer, M. Coppey and B. Hajj for enlightening discussion on the PALM technique, and S. Tlili for comments on the manuscript.

[1] C. W. McCutchen, Journal of the Optical Society of America **57**, 1190 (1967).
[2] A. J. Koster and J. Klumperman, Nature reviews. Molecular cell biology **Suppl**, SS6 (2003).
[3] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, Science **313**, 1642 (2006).
[4] Z. Liu, L. D. Lavis, and E. Betzig, Molecular cell **58**, 644 (2015).
[5] L. Schermelleh, R. Heintzmann, and H. Leonhardt, The Journal of Cell Biology **190**, 165 (2010).
[6] E. Betzig, Optics Letters **20**, 237 (1995).
[7] H. Wolfenson, G. Meacci, S. Liu, M. R. Stachowiak, T. Iskratsch, S. Ghassemi, P. Roca-Cusachs, B. OShaughnessy, J. Hone, and M. P. Sheetz, Nature Cell Biology (2015), 10.1038/ncb3277.
[8] C. Bertocchi, W. I. Goh, Z. Zhang, and P. Kanchanawong, Critical Reviews in Biomedical Engineering **41**, 281 (2013).
[9] R. Changede, X. Xu, F. Margadant, and M. P. Sheetz, Developmental Cell , 1 (2015).
[10] K. H. Biswas, K. L. Hartman, C.-H. Yu, O. J. Harrison, H. Song, A. W. Smith, W. Y. C. Huang, W.-C. Lin, Z. Guo, A. Padmanabhan, S. M. Troyanovsky, M. L. Dustin, L. Shapiro, B. Honig, R. Zaidel-Bar, and J. T. Groves, Proceedings of the National Academy of Sciences of the United States of America **112**, 10932 (2015).
[11] G. Shtengel, Y. Wang, Z. Zhang, W. I. Goh, H. F. Hess, and P. Kanchanawong, *Methods in Cell Biology*, 1st ed., Vol. 123 (Elsevier Inc., 2014) pp. 273–294.
[12] I. Schoen, J. Ries, E. Klotzsch, H. Ewers, and V. Vogel, Nano letters **11**, 4008 (2011).
[13] K. I. Willig, R. R. Kellner, R. Medda, B. Hein, S. Jakobs, and S. W. Hell, Nature Methods **3**, 721 (2006).
[14] P. Hall, *Introduction to the theory of coverage processes* (John Wiley & Sons Australia, Limited, 1988) p. 408.
[15] C. E. Shannon, Proceedings of the IEEE **86**, 447 (1998).
[16] H. Shroff, C. G. Galbraith, J. a. Galbraith, and E. Betzig, Nature methods **5**, 417 (2008).
[17] S. van de Linde, A. Löschberger, T. Klein, M. Heidbreder, S. Wolter, M. Heilemann, and M. Sauer, Nature protocols **6**, 991 (2011).
[18] I. Izeddin, C. G. Specht, M. Lelek, X. Darzacq, A. Triller, C. Zimmer, and M. Dahan, PLoS ONE **6**, e15611 (2011).
[19] U. Endesfelder, S. van de Linde, S. Wolter, M. Sauer, and M. Heilemann, ChemPhysChem **11**, 836 (2010).
[20] B.-C. Chen, W. R. Legant, K. Wang, L. Shao, D. E. Milkie, M. W. Davidson, C. Janetopoulos, X. S. Wu, J. A. Hammer, Z. Liu, B. P. English, Y. Mimori-Kiyosue, D. P. Romero, A. T. Ritter, J. Lippincott-Schwartz, L. Fritz-Laylin, R. D. Mullins, D. M. Mitchell, A.-C. Reymann, J. N. Bembenek, R. Bohme, S. W. Grill, J. T. Wang, G. Seydoux, U. S. Tulu, D. P. Kiehart, E. Betzig, A.-C. Reymann, R. Bohme, S. W. Grill, J. T. Wang, G. Seydoux, U. S. Tulu, D. P. Kiehart, and E. Betzig, Science

346, 1257998 (2014), arXiv:NIHMS150003 .
[21] S. Cox, E. Rosten, J. Monypenny, T. Jovanovic-Talisman, D. T. Burnette, J. Lippincott-Schwartz, G. E. Jones, and R. Heintzmann, Nature Methods **9**, 195 (2011).
[22] A. Triller and D. Choquet, Trends in Neurosciences **28**, 133 (2005).
[23] M. F. Juette, T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf, Nature Methods **5**, 527 (2008).
[24] G. Shtengel, J. A. Galbraith, C. G. Galbraith, J. Lippincott-schwartz, J. M. Gillette, S. Manley, R. Sougrat, C. M. Waterman, P. Kanchanawong, M. W. Davidson, R. D. Fetter, and H. F. Hess, (2008).
[25] B. Hajj, J. Wisniewski, M. El Beheiry, J. J. Chen, A. Revyakin, C. Wu, and M. Dahan, Proc Natl Acad Sci U S A **111**, 17480 (2014).
[26] R. Galland, G. Grenci, A. Aravind, V. Viasnoff, V. Studer, and J.-B. Sibarita, Nature methods **12**, 641 (2015).
[27] A. Martinez-Marrades, J.-F. Rupprecht, M. Gross, and G. Tessier, Optics express **22**, 29191 (2014).
[28] P. Annibale, S. Vanni, M. Scarselli, U. Rothlisberger, and A. Radenovic, Nature Methods **8**, 527 (2011).
[29] P. Erdos and A. Rényi, Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei **6**, 215 (1961).
[30] W. Feller, *Wiley Series*, Vol. 2 (1968) p. 509.
[31] R. P. Stanley and H. S. Wilf, *The American Mathematical Monthly*, Vol. 97 (1990) p. 864.
[32] A. Orlitsky, N. P. Santhanam, and J. Zhang, Science (New York, N.Y.) **302**, 427 (2003).
[33] S. Cox, E. Rosten, J. Monypenny, T. Jovanovic-Talisman, D. T. Burnette, J. Lippincott-Schwartz, G. E. Jones, and R. Heintzmann, Nature Methods **9**, 195 (2011).
[34] D. J. Newman, Science (New York, N.Y.) **98**, 104 (1943).
[35] L. Flatto, Israel J. Math. **15**, 167 (1973).
[36] H. Solomon, *Geometric Probability* (SIAM, 1978) p. 174.
[37] H. Zhang and J. C. Hou, The Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing , 121 (2004).
[38] G. L. Lan, Z. M. Ma, and S. S. Sun, "Discrete Geometry, Combinatorics and Graph Theory: 7th China-Japan Conference, CJCDGCGT 2005, Tianjin, China, November 18-20, 2005, Xi'an, China, November 22-24, 2005, Revised Selected Papers," (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007) Chap. Coverage P, pp. 88–100.
[39] M. I. Stockman, Physics Today **64**, 39 (2011).
[40] H. Cang, A. Labno, C. Lu, X. Yin, M. Liu, C. Gladden, Y. Liu, and X. Zhang, Nature **469**, 385 (2011).
[41] B. Efron, The Annals of Statistics **7**, 1 (1979), arXiv:arXiv:1306.3979v1 .
[42] I. M. Ryzhik and I. S. Gradstein, *Tables of Series, Products and Integrals* (1957) p. 438.
[43] H. Shroff, C. G. Galbraith, J. A. Galbraith, H. White, J. Gillette, S. Olenych, M. W. Davidson, and E. Betzig, Proceedings of the National Academy of Sciences of the United States of America **104**, 20308 (2007).
[44] V. Rivoirard and G. Stoltz, *Statistique mathématique en action* (Vuibert, 2012) p. 448.
[45] In the mathematical literature, this random process is called a Poisson Point Process. The corresponding process where each points are attributed a finite extension σ is called a Boolean Process (see [14]).

Appendix A: Image completion time within the box-filling framework

1. Maximum likelihood estimator of the number of pixel in the PALM

We consider the result of a particular simulation or experiment in which the cumulative number of observation within the pixel j reads: $M_j^{(t)} = k_j$ for all $j \leq F$. The likelihood of such outcome is defined as the probability:

$$\mathbb{P}\left(M^{(t)} = k\right) = \prod_{j=1}^{F} \binom{t}{k_j} p_1^{k_j} (1-p_1)^{t-k_j} 1_{\widehat{F}^{(t)} \leq F}, \quad \text{(A1)}$$

for all $k \geq 0$, and where $1_{\widehat{F}^{(t)} \leq F}$ is the indicative function, equal to 1 if $\widehat{F}^{(t)} \leq F$ and 0 otherwise. The product in Eq. (A1) spans from $j = 1$ to $j = F$ as $M_j^{(t)} = 0$ with probability 1 for all pixels which do not correspond to the structure of interest. Due to the indicative function, the global minimum of Eq. (A1) is achieved for $F = \widehat{F}^{(t)}$ – therefore $\widehat{F}^{(t)}$ is called the maximum likelihood estimator of F.

2. Multi-colored images

Our results are readily adaptable to the case of a colored image, i.e. resulting from the combination of several channels of light emission produced by different imaging probes. This technique is frequently used in cell biology to image simultaneously actin, myosin and other proteins [43]. The number of distinct types of imaging probes is denoted C ; the number of pixels that contain the j–type probe is denoted F_j ; the probability (per pixel and per frame) to detect an imaging probe is denoted $p_{1,j}$. The estimators $\widehat{F}_j^{(t)}$ are defined similarly to Eq. (3). The imaging time is now defined by the relation $\mathbb{P}\left(\left\{\widehat{F}_1^{(t)} = F_1\right\} \cap \ldots \cap \left\{\widehat{F}_C^{(t)} = F_C\right\}\right) = 1 - \theta$. Assuming that the imaging probes act independently, the imaging time t_θ (with C colors) is given by the relation:

$$\prod_{j=1}^{C} \left(1 - (1-p_{1,j})^{t_\theta}\right)^{F_j} = 1 - \theta. \quad \text{(A2)}$$

The imaging time defined in Eq. (A2) exhibit a coupon collector scaling in the following two situations:

(i) if the emission probabilities are identical for all probes (i.e. $p_{1,j} = p_1$), the expression from Eq. (7) holds after the substitution of (a) F by $F_1 + \ldots + F_C$, and (b) μ by $\mu = \mu_1 + \ldots + \mu_C$, which corresponds to the total number of events per frame. Therefore, if we further assume that $F_j = F$ for all j, we show that the

imaging time exhibits a coupon-collector type behavior ($CF\ln(CF)$) in terms of total number of pixels CF.

(ii) if one channel is characterized by a weak blinking probability compared to all the other probes (e.g. $p_{1,1} \ll p_{1,j}$ for all other j, then it will likely be the limiting factor in the imaging process, in which case Eq. (7) holds after the substitution of F and p_1 by F_1 and $p_{1,1}$.

3. Proofs for the coupon-collector scaling in the regime $1 < r < \ln(F)$

a. *One observation per pixel per frame*

Here we consider the case where the number of observations of a pixel at each frame is either 0 or 1 (i.e. $p_k = 0$ for all $k > 1$).

We define the probability $q_j^{(t)}$ that the pixel m has been observed a number j times at the time t: $q_j^{(t)} = \mathbb{P}(M_m^{(t)} = j)$. Successive observations are considered as independent in time, hence $q_j^{(t+1)} = p_0 q_j^{(t)} + p_1\, q_{j-1}^{(t)}$, $\quad 1 \le j \le r-1$ for all $1 \le j < r$. As we are interested in the time required to reach the state $j = r$, we consider the state $j = r$ to be an absorbing state $q_r^{(t+1)} = q_r^{(t)} + p_1\, q_{r-1}^{(t)}$. As soon as $j \le t$, the probability to have reached $j \le r-1$ observations of the pixel is:

$$\mathbb{P}(M_m^{(t)} = j) = \frac{t!}{(t-j)!\, j!}\, p_1^j\, (1-p_1)^{t-j}, \qquad j \le r-1, \tag{A3}$$

from which we deduce the probability that the pixel has been observed at least $r \ge 2$ times is: $q_r^{(t)} = 1 - \sum_{j=0}^{r-1} q_j^{(t)}$. In the long-time limit $1 \ll t$, $t!/(t-j)! \sim t^j$ and the absorption probability $q_r^{(t)}$ tends to 1 as

$$\mathbb{P}(M_m^{(t)} = r) = 1 - t^{r-1} \frac{p_0^t}{(r-1)!} \left(\frac{p_1}{p_0}\right)^{r-1} \quad \text{for} \quad 1 \ll t. \tag{A4}$$

The probability that all pixels have been observed r times at the time t is $(q_r^{(t)})^F$. We are interested in the centile time t_θ given by the condition: $P\left(\{F\delta_{jr}\}_j\right) = (q_r^{(t_\theta)})^F = 1-\theta$. In order to obtain a simple explicit expression for t_θ, we approximate the probability $q_r^{(t)}$ by its long-time behavior from Eq. (A4) (which is valid for θ is sufficiently small or for F sufficiently large) to obtain that:

$$1 - (1-\theta)^{1/F} = \frac{p_0^{t_\theta}}{(r-1)!} \left(t_\theta \frac{p_1}{p_0}\right)^{r-1}. \tag{A5}$$

Given that $1 - (1-\theta)^{1/F} \sim \theta/F$ in the limit $\theta \ll 1$, we obtain from Eq. (A5):

$$\ln(p_0) t_\theta + (r-1)\ln\left(\frac{p_1}{p_0} t_\theta\right) = \ln\left(\frac{\theta}{F}\right) + \ln\left[(r-1)!\right], \tag{A6}$$

which, in the limit $r \ll F$, leads to:

$$t_\theta = -\frac{\left\{\ln\left(\frac{F}{\theta}\right) + (r-1)\ln\left[\frac{p_1}{p_0(-\ln(p_0))}\ln\left(\frac{F}{\theta}\right)\right] + C_1\right\}}{\ln(p_0)}, \tag{A7}$$

where C_1 is a constant of F. In the regime of rare hits ($1 - p_0 \ll 1$), then $\ln(p_0) = \ln(1-(1-p_0)) = -(1-p_0) = -p_1 = -\mu/F$, where μ is the mean number of hits per frame, Eq. (10) then reads

$$t_\theta \underset{1 \ll F}{\sim} \frac{F}{\mu}\left\{\ln(F/\theta) + (r-1)\ln(\ln(F/\theta)) + C_1/F\right\}, \tag{A8}$$

in the limit $r \ll F$. This proves the relation of Eq. (8).

b. *Random number of observations per frame*

In this section, we consider that, at each frame, the number of observations of a given pixel is random variable equal to (i) 0 with probability p_0 and (ii) to a value $k \in [0, r]$ with a probability law p_k.

Following the method of the previous paragraph A 3 a, we consider the coverage dynamic for a single pixel. The probability that the single pixel has been observed j-times, with $1 \le j \le r-1$, during a sequence of t frames is:

$$q_j^{(t)} = \sum \frac{t!}{(t-j_u)!\ldots j_r!} p_0^t \left(\frac{p_1}{p_0}\right)^{j_1} \ldots \left(\frac{p_r}{p_0}\right)^{j_r}, \tag{A9}$$

where (i) the sum holds over the sets of indices $(j_1, \ldots j_r)$ that guarantee the condition that $\sum_{m=1}^r m j_m = j$, and (ii) $j_u = \sum_{m=1}^r j_m$ is the total number of adsorption events.

The imaging time t_θ is defined by the equation: $(q_r^{(t_\theta)})^F = 1-\theta$, in agreement with Eq. (4). In order to obtain a more explicit expression for t_θ, we focus on two simple cases where the asymptotic behavior of $q_r^{(t)} = 1 - \sum_{j=0}^{r-1} q_j^{(t)}$ for $t \gg 1$ can be analytically studied.

We first review the case where steps are all of equal height: $p_k = p_s \delta(s-k)$ (e.g. $p_1 = 0$) and r is a multiple of s i.e. there exists q such that $r = qs$. Then the situations amounts to the case considered in the section A 3 a, with the substitution $r \leftarrow q$ and $p_1 \leftarrow p_s$.

The second case relies on the hypothesis that $p_1 > 0$. The set of indexes that maximizes the exponent j_u in Eq. (A9) under the constraint that $\sum_{m=1}^r m j_m = j$ is

$(j, 0, \dots 0)$. Moreover, $j = r-1$ maximizes the exponent $j_u = j_1 = j$. At the leading order in t, Eq. (A9) reads

$$\mathbb{P}(M_m^{(t)} = r) = 1 - t^{r-1}\frac{p_0^t}{(r-1)!}\left(\frac{p_1}{p_0}\right)^{r-1}, \qquad \text{(A10)}$$

which is identical to Eq. (A4), and leads to the scaling Eq. (A8), and which therefore proves Eq. (10).

Appendix B: Patch-method

In the following sections Sec. B 1 to Sec. B 2, we focus on the 1D–coverage problem of a circle by circular arcs.

1. Evolution of the mean coverage

We denote by $\widehat{Y}^{(t)}$ the fraction of points which are still left uncovered after t frames. For a single patch, $F = 1$, the mean uncovered area is $\mathbb{E}\left[\widehat{Y}|1\,\text{events}\right] = 1 - \sigma/S$. Since patches occur at independent positions within S, we can factorize the mean covered area after n number of patches: $\mathbb{E}\left[\widehat{Y}|n\,\text{events}\right] = (1 - s/S)^n$. After averaging over the distribution of the number of events up to the time t, we find that the mean covered volume after t frames reads [36]

$$\mathbb{E}\left[\widehat{S}^{(t)}/S\right] = 1 - \exp(-\mu t\sigma/S), \qquad \text{(B1)}$$

which holds for arbitrary values of σ/S and t.

2. Evolution of the probability distribution as a function of the number of frames

Following [36], we express the exact probability that n patches cover the whole circle as

$$\mathbb{P}_S\left[0\,\text{gap}|n\,\text{events}\right] = \sum_{j=0}^{k}\binom{n}{j}(-1)^j\left(1 - j\frac{\sigma}{S}\right)^{n-1}, \qquad \text{(B2)}$$

where k is the greatest integer smaller than S/σ, and $\mathbb{P}_S$ denotes the probability measure with a structure volume equal to S.

The probability that no gap remains after a time t then reads: $P\left[0\,\text{gaps}\right] = \sum_{n=k+1}^{\infty} P\left[0\,\text{gaps}|n\,\text{events}\right] P\left[N_e^{(t)} = n\right]$. We define the set of coefficients that

$$a_j^{(t)} = \sum_{n=k+1}^{\infty}\frac{\exp(-\mu t)\left(\binom{n}{j}((\mu t)(1 - j\sigma/S))^n\right)}{n!} = (\mu t)^j e^{-j(\mu t)\sigma/S}(1 - j\sigma/S)^j\gamma_j,$$

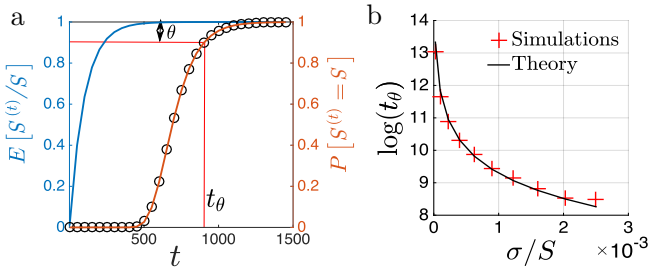

Figure 6: (Color online) (a) Probability of the full coverage of a circle of circumference S by a number of t of patches (arcs segments) of size $\sigma/S = 10^{-2}$: (blue solid line) mean covered area $\widehat{Y}^{(t)}$ after t-frames (with 1 event per frame) ; (orange line) exact probability of a full coverage at time t from Eq. (B2) (black circles) estimated probability for a full coverage at time t, as constructed from the mean covered area $\widehat{Y}^{(t)}$, from Eq. (B5). (b) ...

where

$$\gamma_j = \frac{\Gamma(-j+k+1) - \Gamma(-j+k+1, (\mu t)(1 - j\sigma/S))}{\Gamma(j+1)\Gamma(-j+k+1)}.$$

With these definitions, we conclude using Fubini's theorem that the probability reads

$$P\left[T \le t\right] = P\left[0\,\text{gap}\right] = \sum_{j=0}^{k}(-1)^j\frac{a_j^{(t)}}{1 - j\sigma/S}. \qquad \text{(B3)}$$

We mention that, from Ref. [36], the probability that a number $i \le k$ of gaps remains after n events, denoted $P\left[i\,\text{gaps}|n\,\text{events}\right]$, reads:

$$\binom{n}{i}\sum_{j=0}^{k-i}\binom{n-i}{j}(-1)^j\left(1 - (i+j)\frac{\sigma}{S}\right)^{t-1}, \qquad \text{(B4)}$$

which is the continuous analogue of Eq. (5) in the box-filling model.

3. The area estimation problem

Following the method presented in the final section in the main text, we propose an estimator of the risk that the image is not complete, based on the collected information after t frames. In particular, we provide an estimate of the additional number of frames that should be taken to obtain a complete image at a given confidence $(\widehat{t}_\theta^{(t)})$. This estimator rely on the current covered area of the experimental realization $\widehat{S}^{(t)}$, which corresponds to a maximum-likelihood estimator of the volume of the structure to be imaged. We assume that the covered length $\widehat{F}^{(t)}$ is well described by the mean covered area at time t, hence that

$$\widehat{\mathbb{P}}^{(t)}\left(\widehat{S}^{(t)} = S\right) \approx \mathbb{P}_{S\left(1 - e^{-\sigma t/S}\right)}\left[0\,\text{gap}\right], \qquad \text{(B5)}$$

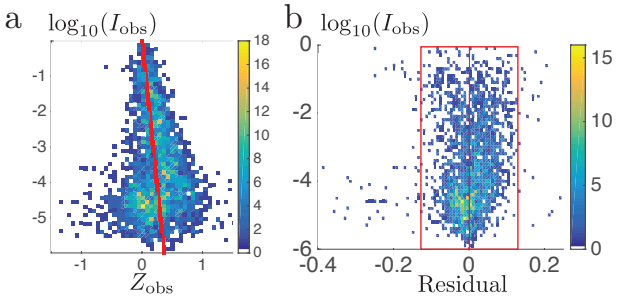

Figure 7: (Color online) Analysis of the experiments from Ref. [27] using the Gaussian model defined in Eq. (C1). (a) Density representation of the experimental data: measured scattered intensity I_i as a function of the measure height Z_i: (colormap) experimental number of observations per boxes (red line) linear regression for $\ln(I)$ in terms of z. (b) Residual analysis and validation the linear Gaussian model: we verify that residual $R_i = \sqrt{I_i}Z_i + \sqrt{I_i}\ln(I_i/\widehat{I}_{0n}))$ is centered (red line) and maintains a constant variance (red box) with the intensity $I_{\rm obs}$.

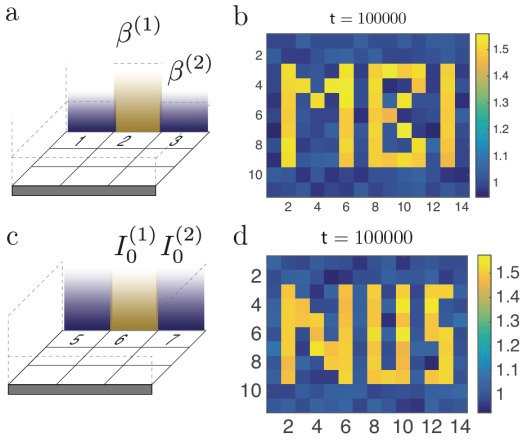

Figure 8: (Color online) Convergence of estimators and the imaging process of a patterned surface (14×14 pixels) in terms of the penetration length field β and maximal intensity I_0. (a) The penetration length is $\beta = \beta^{(2)}$ by default and $\beta = \beta^{(1)} = 1.50\beta^{(2)}$ within specific pixels (MBI pattern). (b) Convergence of the estimator $\widehat{\beta}_n$, defined in Eq. (C2). (c) The maximal intensity $I_0 = I_0^{(1)}$ or $I_0 = I_0^{(2)} = 0.75 I_0^{(1)}$ within specific pixels (NUS pattern). (d) Convergence of the estimator $\widehat{I}_{0n}$ defined in Eq. (C3).

where $\mathbb{P}_{S\left(1-e^{-\mu\sigma t/S}\right)}$ refers to the probability defined in Eq. (B2), that the total area to be covered is $S\left(1-e^{-\mu\sigma t/S}\right)$. We test the method on Fig. SI. 6a, in a case where the ratio S/σ is known, and we find that the approximation of Eq. (B5) is very satisfactory. Following the procedure defined in Sec. , we estimate the additional number of frames required to obtained a complete image, which can be deduced from Eq. (9), through the substitution of S and μ by their corresponding estimators.

Appendix C: Calculation of the minimum number r of observations per pixel

In this section, we justify that a large number of observations is required per pixel to obtain a reliable measurements of the electromagnetic field considered in Ref. [27]. We recall that the experimental data consists in 2792 measurements of heights and intensities (Z_i, I_i) $(i \leq 2792)$. We model the noise on the height measurement through the following linear Gaussian model:

$$I_i Z_i = -\beta\sqrt{I_i}\ln\left(I_i/I_0\right) + \sigma\eta_i, \tag{C1}$$

where η is a standard Gaussian white noise process. We define the vector of unknown parameters $B = \beta(-1, \ln I_0)$, where β refers to the penetration length and I_0 to the intensity at the surface, as well as $Y = \sqrt{I_i}Z_i$ (vector of observations) and $X = (\sqrt{I_i}\ln I_i, \sqrt{I_i})$ (explicative matrix). In terms of these quantities, the model defined in Eq. (C1) reads $Y = XB + \sigma\eta$, which corresponds to the well-known linear Gaussian model. The estimator of the vector B that maximizes the likelihood function is $\widehat{B}_n = (X^T X)^{-1} X^T Y$ [44] ; the developed expression corresponds to the two estimators

$$\widehat{\beta}_n = -C\left\{\overline{I}\cdot\overline{ZI\ln I} - \overline{I\ln I}\cdot\overline{ZI}\right\}, \tag{C2}$$

$$\widehat{I}_{0n} = \exp\left(\frac{1}{\widehat{\beta}_n}\left\{\overline{I}\cdot\overline{ZI\ln I} - \overline{I\ln I}\cdot\overline{ZI}\right\}\right), \tag{C3}$$

where $\overline{I} = \sum_{i=1}^n I_i$ and $C = \left(\overline{I}\cdot\overline{I\ln^2 I} - \overline{I\ln I}^2\right)^{-1}$. The variance of the noise is also an unknown variable that can be evaluated by the estimator:

$$\widehat{\sigma}_n^2 = (\sum_{i=1}^n R_i^2)/(n-2), \tag{C4}$$

where $R_i = \sqrt{I_i}Z_i + \sqrt{I_i}\ln(I_i/\widehat{I}_{0n}))$ is called the residual. For a Gaussian distribution of noise, we expect to have: $\widehat{\sigma}_n^2/\sigma^2 \sim \mathcal{F}\left(1, 1/\sqrt{n-2}\right)$ in the limit $n \gg 1$.

We now define confidence intervals for the estimators defined in Eqs. (C2–C3). We consider a risk level $\alpha = 0.05$: with a $1-\alpha = 0.95$ probability, the quantity β lies within the confidence interval $\mathcal{C}(\widehat{\beta}_n)$:

$$\beta \in \mathcal{C}(\widehat{\beta}_n) = \left[\widehat{\beta}_n \pm \widehat{\sigma}_n t_{1-\alpha/2}^{(n-2)}\sqrt{\overline{I}C}\right], \tag{C5}$$

where $t_{0.975}^{(n-2)}$ is the one-sided quantile of the Student distribution, with $t_{0.975}^{(n-2)} = 1.96\ldots$ in the limit $n \gg 1$. As the quantity C is inversely proportional to the number of observations n, the confidence interval Eq. (C5) narrows on β with a $1/\sqrt{n}$ speed as n increases.

From Eq. (C5), we obtain an estimate of r, i.e. the minimal number $n = r$ of observations required so that

the estimator of β has an error lesser than 10%, with probability 95% probability. From experimental values, we find that the confidence interval for β is $60,1 \pm 7.3\,\mathrm{nm}$ for $n = 2792$.

Therefore, the estimate for the minimal observation per pixel r should around 4000.

We now consider a system of F identical pixels. The probability that all the estimators $\widehat{\beta}_n^{(j)}$, where $1 \leq j \leq F$ is the pixel label, are precise at 10% to the exact value β is $\mathbb{P}\left[\forall j, \beta \in CI(\widehat{\beta}_n^{(j)})\right] = (\mathbb{P}\left[\beta \in CI(\widehat{\beta}_n^{(1)})\right])^F$. We set $\mathbb{P}\left[\forall j, \beta \in CI(\widehat{\beta}_n^{(j)})\right] = 1 - \epsilon$ where ϵ is the accepted risk level (in the following $\epsilon = 0.05$). Therefore, compared to the case of single pixel, the risk level on a single pixel is to be divided by a factor F: $\mathbb{P}\left[\beta \in CI(\widehat{\beta}_n^{(1)})\right] = 1 - 0.05/F$. The confidence interval from Eq. (C5) holds provided that $\alpha = 0.05/F$, which corresponds to a narrower confidence interval compared to the single pixel case. However the quantity $t_{1-\epsilon/(2F)}^{(n-2)}$ increases weakly with the number of pixels F (e.g. $t_{1-0.05/(2\times 10)}^{(\infty)} \approx 2.80$ and $t_{1-0.05/(2\times 100)}^{(\infty)} \approx 3.40$).

In conclusion, increasing the number of pixels F does not significantly increase the required number of observations per pixel r.

We now apply our results to the detection of a pattern in the electromagnetic field (see Fig. SI. 8), in a simulated experiment. We consider a local electromagnetic field that takes the form of Eq. (2), and in which β and I_0 may take either of two values. In Fig. SI. 8, the error of the estimators β is lower than 10% ; the image indeed results from the superposition of $t = 10^6$ frames, a number in agreement with the predicted threshold from Eq. (12): $F \times r \approx 10^6$, where $F = 14 \times 14$ and $r = 4000$.